

# ANÁLISIS DE PUNTO DE CAMBIO

## DENTRO DEL CONTROL ESTADÍSTICO DE PROCESOS

Ángel Salvador Pérez Blanco  
Álvaro Eduardo Cordero Franco  
Victor Tercero Gómez

María Aurora Chávez Valdez  
UANL-FCFM

Universidad Autónoma de Nuevo León  
Facultad de Ciencias Físico Matemáticas  
San Nicolás de los Garza, Nuevo León, México

### Resumen:

El Análisis de Punto de Cambio (CPA *Change Point Analysis* por sus siglas en Inglés) es una parte del Control Estadístico de Calidad (SPC *Statistical Control Process*, por sus siglas en Inglés) que aparece por primera vez alrededor de 1954 y en las últimas tres décadas ha sido ampliamente estudiado por investigadores e ingenieros de diferentes áreas del conocimiento. En este artículo proveeremos una explicación que cubre: sus orígenes, las aportaciones más importantes, las áreas en que se ha aplicado, las líneas actuales de investigación, el futuro y un ejemplo basado en el caso de la función de distribución Gamma. El objetivo principal del presente artículo es familiarizar a los lectores con esta nueva metodología.

### Palabras claves:

punto de cambio, procesos Gamma, control estadístico de procesos, estimador de máxima verosimilitud

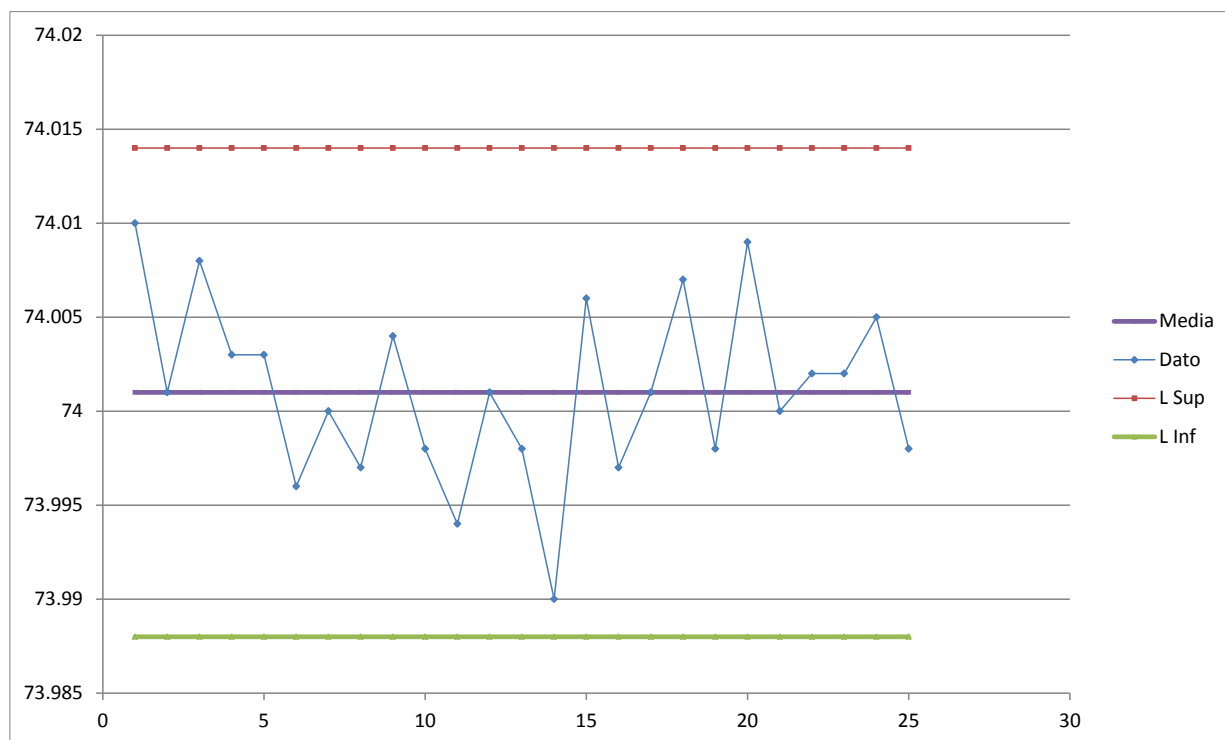
## Introducción

El Control Estadístico de Procesos (SPC *Statistical Process Control*) busca mantener un proceso en un estado de control estadístico, lo cual implica que sea predecible para permitir su manejo. Uno de los pioneros en esta área del conocimiento y que hizo aportaciones invaluable es Shewhart [1]; considerado como uno de los principales estudiosos del SPC (*Statistical Process Control*) y también de los que más ha aportado al mismo, el libro que publicó en 1931 sigue siendo motivo de estudio y las gráficas de control que propuso siguen siendo un estándar en la industria y la investigación. Las gráficas de control fueron de las primeras herramientas desarrolladas dentro del SPC y son de utilidad para monitorear sistemas, sus fundamentos teóricos y prácticos, iniciales, parten de suponer que la característica que se está midiendo de las muestras de un proceso siguen algún tipo de función de distribución de probabilidad, si bien esta limitación inicial está siendo superada, los esfuerzos para lograrlo datan de los últimos 10 años. Un caso de estos esfuerzos puede ser consultado en Tercero et al [2], donde se presentan herramientas que permiten monitorear un proceso en el que las suposiciones de normalidad no son asumidas. Una gráfica de control implica un valor medio, un límite de control superior y un límite de control inferior (UCL *Upper Control Limit*, LCL *Lower Control Limit*, respectivamente), mientras el valor medido de una característica de calidad del producto está dentro de los límites de control se dice que el proceso se encuentra en control estadístico, cuando dicha característica medida está fuera de los límites de control establecidos, se considera que algo perturbó al proceso y, en consecuencia, los parámetros del modelo que lo definen han cambiado. Se dice que el proceso está fuera de control estadístico. El momento en que el cambio ocurre es llamado el Punto de Cambio (CP *Change Point*) y el Análisis de Punto de Cambio (CPA *Change Point Analysis*) es la herramienta que estudia este fenómeno. Debido a que el punto de cambio es desconocido el CPA se dedica a estimar este valor, ya que esta estimación será utilizada para determinar las causas posibles que han provocado que el sistema esté fuera de control estadístico. Una vez que se ha determinado el posible momento en que el proceso empezó a tener problemas es posible reducir el área de búsqueda de la causa que ha generado este disturbio y, una vez identificada, tomar una acción correctiva que permita volver a tener al proceso en control estadístico. Algunas preguntas que surgen cuando una gráfica de control detecta que el proceso ha salido de control son: ¿Qué ocurrió con las observaciones previas a la observación

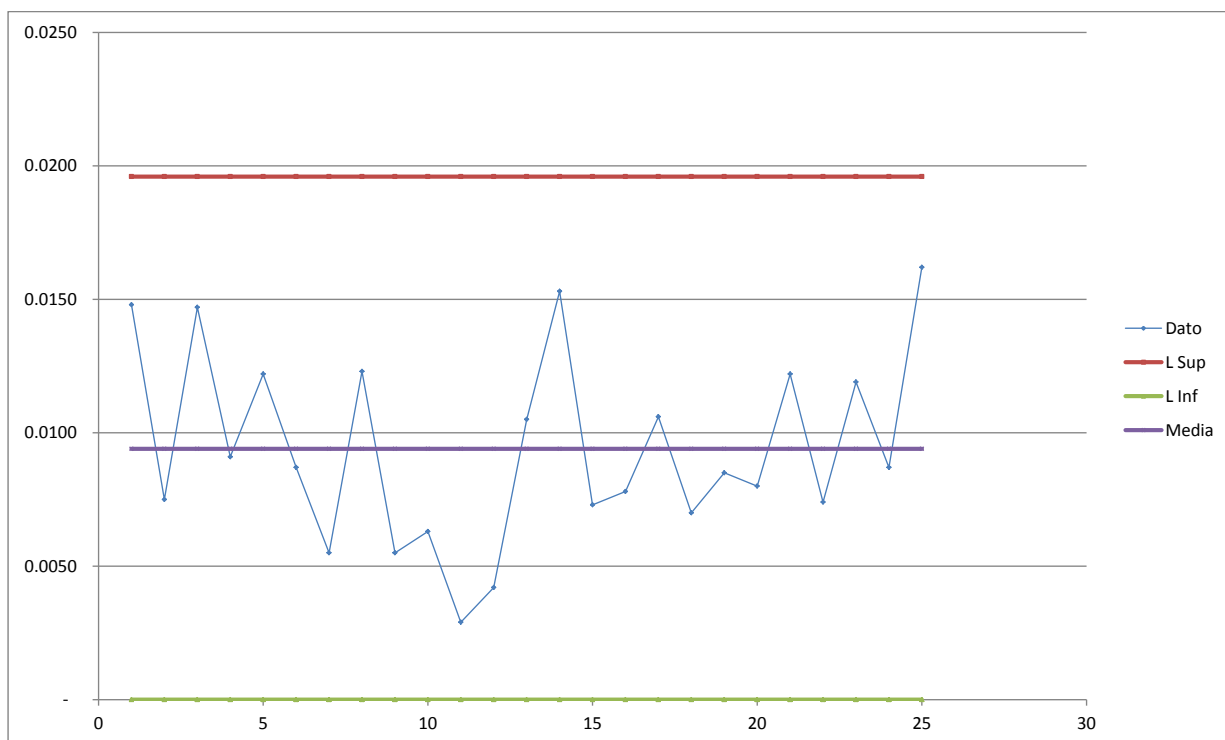
que se encontró fuera de control estadístico? ¿También estaban fuera de control pero no fueron detectadas a tiempo? ¿De esa serie de observaciones cuál es la primera donde el proceso ha cambiado? Los autores esperamos que después de la lectura de este artículo los lectores se aventuren a proponer sus propias respuestas y consideren confrontarlas con lo que ya está escrito al respecto. Con el fin de aterrizar los conceptos se ha preparado un ejemplo para la estimación del punto de cambio basado en un proceso que considera un conjunto de datos obtenidos como una serie en el tiempo y que pueden ser modelados mediante una función de distribución Gamma. Se ha seleccionado la función de distribución Gamma por ser utilizada para modelar procesos en que el tiempo entre llegadas es relevante, y debido a sus aplicaciones en diferentes áreas tal como la industria química, el control de inventarios y los estudios biológicos [3] [4], además de ser una distribución que no ha sido tan estudiada, bajo la óptica del SPC (*Statistical Process Control*), como lo ha sido la distribución Normal. Para un conjunto de datos particular se mostrará como calcular la gráfica de control EWMA (*Exponentially Weighted Moving Average*) y como estimar el punto de cambio mediante la técnica del MLE (*Maximum Likelihood Estimation*). En las Figuras 1 y 2 se presentan ejemplos de gráficas de control. Como comentario general a las gráficas de control de Shewhart, CUSUM y EWA se puede señalar que las gráficas de Shewhart detectan cambios abruptos pero no son eficientes en la detección de cambios lentos y sostenidos, en tal caso las de CUSUM y EWMA funcionan mejor.

## Trabajos previos

Los primeros trabajos sobre CPA (*Change Point Analysis*), dentro del marco del SPC, fueron realizados por Girshick y Rubin [5] utilizando un enfoque Bayesiano. Posteriormente, Page [6] [7] [8] propuso la técnica de CUSUM (*Cummulative SUMs*) en 1954 para detectar cambios sostenidos en series de tiempo. Respecto a procedimientos paramétricos, Hinkley [9] consolidó la teoría correspondiente para encontrar el punto de cambio utilizando MLE (*Maximum Likelihood Estimation*) y pruebas de radio de verosimilitud (LRT *Likelihood Ratio Test*, por sus siglas en inglés) considerando los parámetros iniciales conocidos y desconocidos. Samuel [10] propuso el uso de un estimador después de que una gráfica de control de Shewhart indicara un cambio en el proceso. Este mismo enfoque fue analizado por Nedumaran, Pignatiello y Calvin [11] y Samuel, Pignatiello y Calvin [12]. Una herramienta diferente para crear gráficas de control basadas en un enfoque de



**Figura 1.** Ejemplo de gráfica de control para la  $\bar{X}$



**Figura 2.** Ejemplo de Gráfica de Control para la desviación S.

CPA (*Change Point Analysis*), usando una analogía del radio de verosimilitud generalizado, fue desarrollada por Hawkins y Zamba [13]. Usando una combinación MLE y gráficas de control para determinar una estimación del punto de cambio Shao y Hou [14] mostraron que una gráfica de control EWMA (*Exponentially Weighted Moving Average*) puede ser una forma relativamente buena para estimar el punto de cambio en un proceso Gamma cuando solo el parámetro de escala cambia. Ellos han combinado una gráfica de control EWMA adaptada para la distribución Gamma y los estimadores de máxima verosimilitud para detectar y estimar puntos de cambio, respectivamente. Sus resultados muestran que el MLE fue un mejor estimador del punto de cambio que el indicado por la gráfica de control EWMA. El presente artículo continúa ese trabajo, pero ahora derivando el estimador correspondiente del CPA (*Change Point Analysis*) para el caso en el parámetro de forma () cambia a partir de un momento específico.

### Gráfica de Control Gamma

El modelo propuesto usa una gráfica de control EWMA (*Exponentially Weighted Moving Average*) para detectar los cambios durante el monitoreo del proceso. Una vez que el cambio es detectado el punto de cambio  $\tau$  es estimado usando el MLE (*Maximum Likelihood Estimation*). Se ha considerado una serie de datos independientes provenientes de una función de distribución Gamma  $X_1, \dots, X_\tau, X_{\tau+1}, \dots, X_T$  como se muestra en (1)

$$X_i \sim \begin{cases} G(a_0, b_0), & 1 \leq i \leq \tau \\ G(a_1, b_0), & \tau < i \leq T \end{cases} \quad (1)$$

$a_0$  y  $b_0$  y  $b_1$  son parámetros conocidos, mientras que  $a_1$  es desconocido, y  $T$  es la última observación de la serie de datos y  $a$  es identificado como el parámetro de la forma y  $b$  es identificado como el parámetro de la escala. Los límites de control para la gráfica de control EWMA (*Exponentially Weighted Moving Average*) son calculados mediante:

$$\begin{aligned} UCL &= \mu_0 + L\sigma_0 \sqrt{\frac{\lambda}{2-\lambda} [1-(1-\lambda)^{2i}]} \\ LCL &= \mu_0 - L\sigma_0 \sqrt{\frac{\lambda}{2-\lambda} [1-(1-\lambda)^{2i}]} \end{aligned} \quad (2)$$

Donde  $\mu_0 = a_0 b_0$  y  $\sigma_0 = \sqrt{a_0 b_0^2}$  se refieren a la media y la desviación estándar del proceso respectivamente. El  $\hat{\tau}_{EWMA}$  es definido como el primer valor fuera de control en una gráfica de control EWMA. Si ya no son tomadas más muestras después de que la gráfica de control indica que se ha dado el primer punto fuera de los límites de control se considera  $T = \hat{\tau}_{EWMA}$ . Un estudio detallado de la gráfica de control EWMA puede ser encontrado en Montgomery [15].

### Estimación del punto de cambio

Para realizar la estimación del punto de cambio se usará la teoría desarrollada por Hinkley [9] y se considerará un proceso caracterizado por una función de distribución Gamma donde los cambios en el parámetro ocurren entre el tiempo  $\tau$  y  $\tau + 1$ , además, se asumirán conocidos los parámetros  $a_0, b_0, b_1$  y se considerará que las observaciones son individuales. Bajo estas consideraciones la función de verosimilitud usada para obtener los MLE (*Maximum Likelihood Estimation*) es (3)

$$L(\tau, a_1, b_1 | x) = \prod_{i=1}^{\tau} f(\tau, a_0, b_0 | x_i) \cdot \prod_{i=\tau+1}^T f(\tau, a_1, b_1 | x_i) \quad (3)$$

Sin embargo, es más fácil trabajar con el logaritmo de la función de verosimilitud (4)

$$\begin{aligned} \ln L &= -\tau \ln \Gamma(a_0) - \tau a_0 \ln(b_0) - (T - \tau) \ln \Gamma(a_1) - (T - \tau) a_1 \ln(b_1) \\ &+ (a_0 - 1) \sum_{i=1}^{\tau} \ln(x_i) - \sum_{i=1}^{\tau} \frac{x_i}{b_0} \\ &+ (a_1 - 1) \sum_{i=\tau+1}^T \ln(x_i) - \sum_{i=\tau+1}^T \frac{x_i}{b_1} \end{aligned} \quad (4)$$

Para obtener el estimador de máxima verosimilitud correspondiente a  $a_1$ , recuerde que  $b_0 = b_1$  es conocido, es necesario resolver la ecuación dada por la primera derivada parcial de (4) con respecto a  $a_1$  e igualarla a cero, obteniendo (5) cuya solución para cada caso es sustituido en (6).

Las ecuaciones obtenidas son:

$$(T - \tau) \frac{\Gamma'(\hat{a}_1)}{\Gamma(\hat{a}_1)} = \sum_{i=\tau+1}^T \ln(x_i) - (T - \tau) \ln(\hat{b}_0) \quad (5)$$

$$\hat{\tau}_{MLE} = \operatorname{argmax} \ln L(\hat{a}_1, \tau) \quad (6)$$

Donde el cociente  $\frac{\Gamma'(\hat{a}_1)}{\Gamma(\hat{a}_1)}$  es conocido como la función DiGamma.

### Ejemplo numérico

En el ejemplo que se presenta a continuación se ha simulado un conjunto de datos; los primeros 15 provenientes de una función de distribución Gamma ( $a_0=1, b_0=1$ ), y los posteriores 6 a partir de una función de distribución Gamma ( $a_1=2, b_1=1$ ). Con esto se simula un cambio en el proceso a partir del dato 16. Los límites de control fueron calculados considerando  $\lambda=0.20$  y  $L=2.692$ , los datos y cálculos para la gráfica de control EWMA (*Exponentially Weighted Moving Average*) y el MLE (*Maximum Likelihood Estimation*) son presentadas en la **Tabla 1**. Como se ve en ésta, la gráfica de control EWMA detecta un valor fuera de los límites de control en la observación 21. El punto de cambio estimado por el MLE está en la observación 15, seis datos antes de la detección de la gráfica de control EWMA.

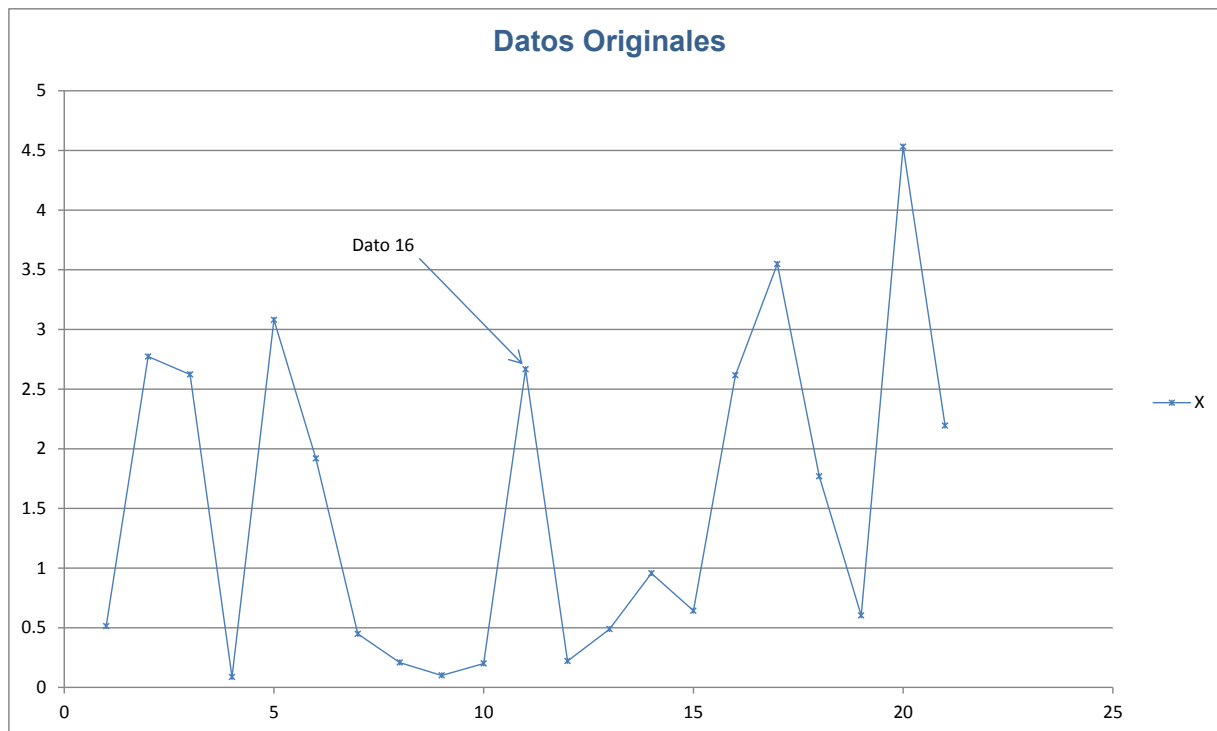
En este ejemplo la observación 16 es el primer valor generado después de que el cambio ocurrió. Las **figuras 3 y 4** muestran el comportamiento de los datos originales y transformados, respectivamente. Para la gráfica de control EWMA (*Exponentially Weighted Moving Average*) los datos originales son transformados utilizando (7).

$$z_i = \lambda x_i + (1-\lambda) z_{i-1} \quad (7)$$

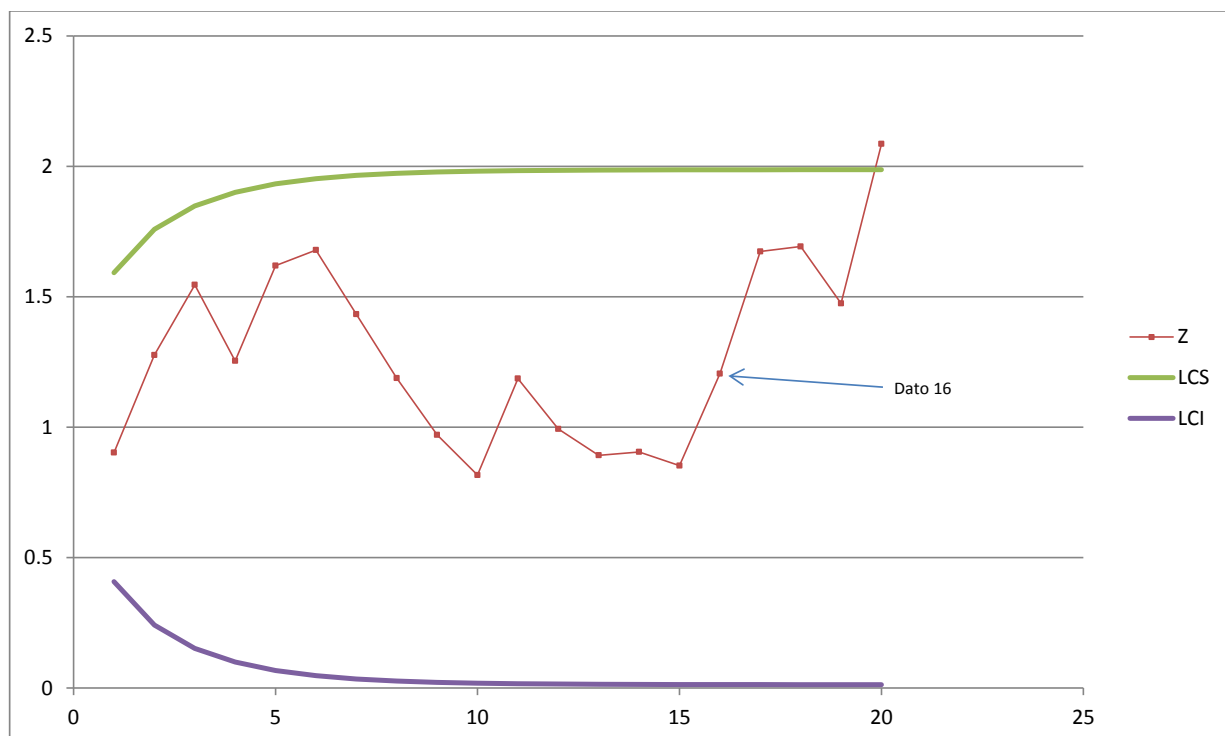
Un elemento interesante en este trabajo es el uso de la función DiGama, la cual consiste en el cociente entre la derivada de la Gamma y la función misma. Esta función es reportada en diferentes aplicaciones y ha sido estudiada por varios autores, siendo Cristinel [16] uno de los autores que maneja formas para calcular el valor.

**Tabla 1** Ejemplo de los cálculos requeridos para calcular el EWMA y el MLE. Los primeros 15 observaciones provienen de una Gamma ( $a_0=1, b_0=1$ ) los siguientes 6 datos provienen de una Gamma ( $a_1=2, b_1=1$ ). Las col. 3-5 corresponden a los cálculos para EWMA y las col. 6 y 7 corresponden a los cálculos para MLE.

	OBSER	EWMA			MLE	
<i>i</i>	$X_i$	$Z_i$	UCL	LCL	$\hat{a}_1$	$\ln(\hat{a}_1 \hat{\tau})$
0		1				
1	0.5129	0.9026	1.5924	0.4076	1.3579	-36.3060
2	2.7727	1.2766	1.7586	0.2414	1.3113	-36.2384
3	2.6224	1.5458	1.8481	0.1519	1.2692	-36.0512
4	0.0869	1.2540	1.9007	0.0993	1.3796	-35.8821
5	3.0802	1.6192	1.9328	0.0672	1.3272	-35.9010
6	1.9186	1.6791	1.9528	0.0472	1.2981	-35.7939
7	0.4483	1.4329	1.9654	0.0346	1.3341	-35.6665
8	0.2083	1.1880	1.9733	0.0267	1.4084	-35.2571
9	0.1002	0.9705	1.9784	0.0216	1.5278	-34.0345
10	0.2001	0.8164	1.9816	0.0184	1.6196	-32.6149
11	2.6657	1.1862	1.9837	0.0163	1.5629	-33.3462
12	0.2207	0.9931	1.9850	0.0150	1.6513	-31.8700
13	0.4883	0.8922	1.9858	0.0142	1.6957	-30.9747
14	0.9561	0.9050	1.9864	0.0136	1.6984	-30.8266
15	0.6424	0.8524	1.9867	0.0133	1.7266	-30.1997
16	2.6155	1.2051	1.9869	0.0131	1.6661	-31.2018
17	3.5481	1.6737	1.9871	0.0129	1.5907	-32.2680
18	1.7685	1.6926	1.9872	0.0128	1.5582	-32.5897
19	0.6028	1.4747	1.9872	0.0128	1.5870	-32.0915
20	4.5316	2.0861	1.9873	0.0127	1.5030	-32.9955
21	2.1936	2.1076	1.9873	0.0127		



**Figura 3.** Gráfica de dispersión de los datos del ejemplo



**Figura 4.** Gráfica de control EWMA para los datos del ejemplo. Los límites han sido calculados usando la ecuación (2).



Es importante resaltar que la ecuación (5) no tiene una expresión explícita por lo que se hace necesario utilizar análisis numérico para encontrar los valores de  $a_1$ .

### Resultados del ejemplo numérico

De la construcción del ejercicio se estableció que el parámetro  $a$  cambiará del valor 1 al 2 a partir del dato 16 inclusive, recuerde que el parámetro  $b$  se mantiene constante en 1. En la fig. 3 no es claro este cambio, sin embargo en la fig. 4 la tendencia de los datos es evidente, este es un resultado de aplicar la gráfica EWMA (*Exponentially Weighted Moving Average*) para la función de distribución Gamma. De las tablas puede apreciarse que el estimador de máxima verosimilitud, en este ejemplo numérico, tiene un sesgo menor que el estimador del EWMA y en consecuencia lo ubica en un valor más cercano al punto de cambio que la gráfica de control EWMA, mientras que esta gráfica manda la señal de alerta en el dato 21, el MLE (*Maximum Likelihood Estimation*) lo hace en el dato 15. Esto es, uno indica el punto de cambio cinco datos después de cuando realmente ocurrió, mientras que el otro lo ubica un dato antes. La eficiencia del método utilizado en este documento puede ser consultada en [17] así como un análisis del caso en que cambian los dos parámetros, al mismo tiempo.

### Conclusiones y trabajos futuros

En este trabajo se ha ofrecido un panorama general de dos conceptos importantes en el SPC (*Statistical Process Control*), las gráficas de control y el análisis de punto de cambio. Estos conceptos han tomado relevancia en los últimos 50 años, presentando un crecimiento muy marcado en las publicaciones especializadas el cual puede ser consultado en [18]. Las gráficas de control enfocadas al CPA (*Change Point Analysis*) han encontrado un campo fértil y continuamente aparecen modificaciones, a las ya conocidas, que permiten hacer una estimación del CP (*Change Point*) con características que mejoran la precisión y/o exactitud y cuyos cálculos pueden o no presentar un gran complejidad. Por su lado el MLE (*Maximum Likelihood Estimation*) se ha integrado a las gráficas de control produciendo estimadores muy cercanos al punto de cambio real. Las aplicaciones resultantes de estas investigaciones y la combinación de las mismas son tan variadas que van desde la detección de terremotos hasta el análisis del comportamiento de crecimiento de epidemias. Desde luego, esto no quiere decir que los procesos industriales esten excluidos o sean los menos, nada más alejado de la realidad.

La finalidad del ejemplo numérico es presentar como se integran ambas herramientas para estimar el punto de cambio. Se han detallado los pasos a seguir con la finalidad de que el lector pueda desarrollar sus propias simulaciones y comprobar los resultados que aquí se presentan así como las bondades de cada herramienta. Si bien es cierto que con un ejemplo no es posible obtener conclusiones generales, también lo es que los artículos consultados para la realización del presente y las aportaciones hechas por los autores del presente respaldan la afirmación sobre la prevalecencia de uno sobre el otro. Es importante señalar que la combinación de las gráficas de control y la estimación del punto de cambio mediante el MLE (*Maximum Likelihood Estimation*) es una herramienta que muestra excelentes resultados. El explicar la forma en que una técnica y otra se complementan de manera natural amerita un artículo por sí mismo.

Finalmente los autores agradecemos a la FCFM de la UANL por el apoyo con que nos ha favorecido para la realización de investigaciones en esta área de conocimiento y tenemos confianza en que la presente aportación motive a los lectores a considerar este enfoque de solución al problema de punto de cambio y se interesen en incursionar en el mismo.

## Referencias

- [1] W. A. Shewhart. Economic Control of Quality of Manufactured Product 50th Anniversary. Asq Press. 1931.
- [2] V. Tercero, M. d. C. Temblador-Pérez, M. Beruvides & A. A. Hernández-Luna. "Nonparametric Estimator for the Time of a Step Change in the Trend of Random Walk Models with Drift". *Quality and Reliability Engineering International*. vol. 28. no. 3. 08 Feb 2012.
- [3] T. Burgin. "The gamma distribution and inventory control". *Operational Research Quarterly*. Vols. 26, 1. no. 3. pp. 507-525. 1975.
- [4] I. Feinberg, H. Thode, H. Chugani & J. March. "Gamma distribution model describes maturational curves for delta wave amplitude, cortical metabolic rate and synaptic density". *Journal of Theoretical Biology*. vol. 142. no. 2. pp. 149-161. 1990.
- [5] M. A. Girshick & H. Rubin. "A Bayes approach to a quality control model". *The Annals of mathematical statistics*. pp. 114-125. 1952.
- [6] E. S. Page. "A test for a change in a parameter occurring at an unknown point". *Biometrika*. vol. 42. no. 3-4. pp. 523-527. 1955.
- [7] E. S. Page. "Continuous inspection schemes". *Biometrika*. vol. 41. no. 1/2. pp. 100-115. Jun 1954.
- [8] E. S. Page. "On problems in which a change in a parameter occurs at an unknown point". *Biometrika*. vol. 44, no. 1-2. pp. 248-252. 1957.
- [9] D. Hinkley. "Inference about the change-point in a sequence of random variables". *Biometrika*. vol. 57. no. 1. pp. 1-17. 1970.
- [10] T. Samuel, J. Pignatiello & J. Calvin. "Identifying the Time of a Step Change with x Control Charts". *Quality Engineering*. vol. 10. no. 3. pp. 521-527. 1998a.
- [11] G. Nedumaran, J. Pignatiello & J. A. Calvin. "Identifying the time of a step-change with  $\chi^2$  control charts". *Quality Control and applied statistics*. vol. 47. no. 2. pp. 125-126. 2002.
- [12] T. R. Samuel, J. J. Pignatiello & J. A. Calvin. "Identifying the time of a step change in Normal process variance". *Quality Engineering*. vol. 10. no. 3. p. 529. 1998.
- [13] D. M. Hawkins & K. D. Zamba. "Statistical process control for shifts in mean or variance using a changepoint formulation". *Technometrics*. vol. 47. no. 2. pp. 164-173. 2005.
- [14] H. C. Shao Y.E. "A combined MLE and EWMA chart approach to estimate the change point of a gamma process with individual observations". *International Journal of Innovative Computing, information and Control*. vol. 7. no. 5A. pp. 2109-2122. 2011.
- [15] D. C. Montgomery. Control Estadístico de la Calidad. México: John Wiley & Sons, Ltd. 2007.
- [16] M. Cristinel. "New approximations of the gamma function in terms of the digamma function". *Applied Mathematics Letters*. vol. 23. no. 1. pp. 97-100. 2010.
- [17] A. Pérez-Blanco, V. Cordero-Franco & V. Tercero-Gómez. "Change Point Estimation After a Gamma Control Chart". in *ISERC 2013*, San Juan, Puerto Rico, 2013.
- [18] A. S. Perez-Blanco, A. E. Cordero-Gómez & V. Tercero-Gómez. "Punto de Cambio: Estado Actual y Tendencias". in *2do Foro de Divulgación Científica y Tecnológica*. Monterrey. 2012.

## Datos de los autores:

**M.C. Ángel Salvador Pérez Blanco**

Email: angel.perezbl@uanl.edu.mx

**Ph.D. Álvaro Eduardo Cordero Franco**

Email: alvaro.corderofr@uanl.edu.mx

**Ph.D. Víctor Tercero Gómez**

Email: victor.tercerog@uanl.edu.mx

**M.C. María Aurora Chávez Valdez**

Email: maria.chavezv@uanl.mx